DOCUMENT RESUME

ED 053 217 TM 000 756

TITLE	1970-71 Individual Pupil Report: Explanatory Materials.
INSTITUTION	Michigan State Dept. of Education, Lansing.
PUB DATE	Apr 71
NOTE	33p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Achievement Tests, Educational Accountability, English, Grade 4, Grade 7, Instructional Materials, Mathematics, Reading Skills, Scores, *State Boards of Education, *State Programs, *Statistical Analysis, *Testing Programs, Test Reliability, Test Validity, Vocabulary Skills, Written Language
IDENTIFIERS	*Michigan Educational Assessment Program

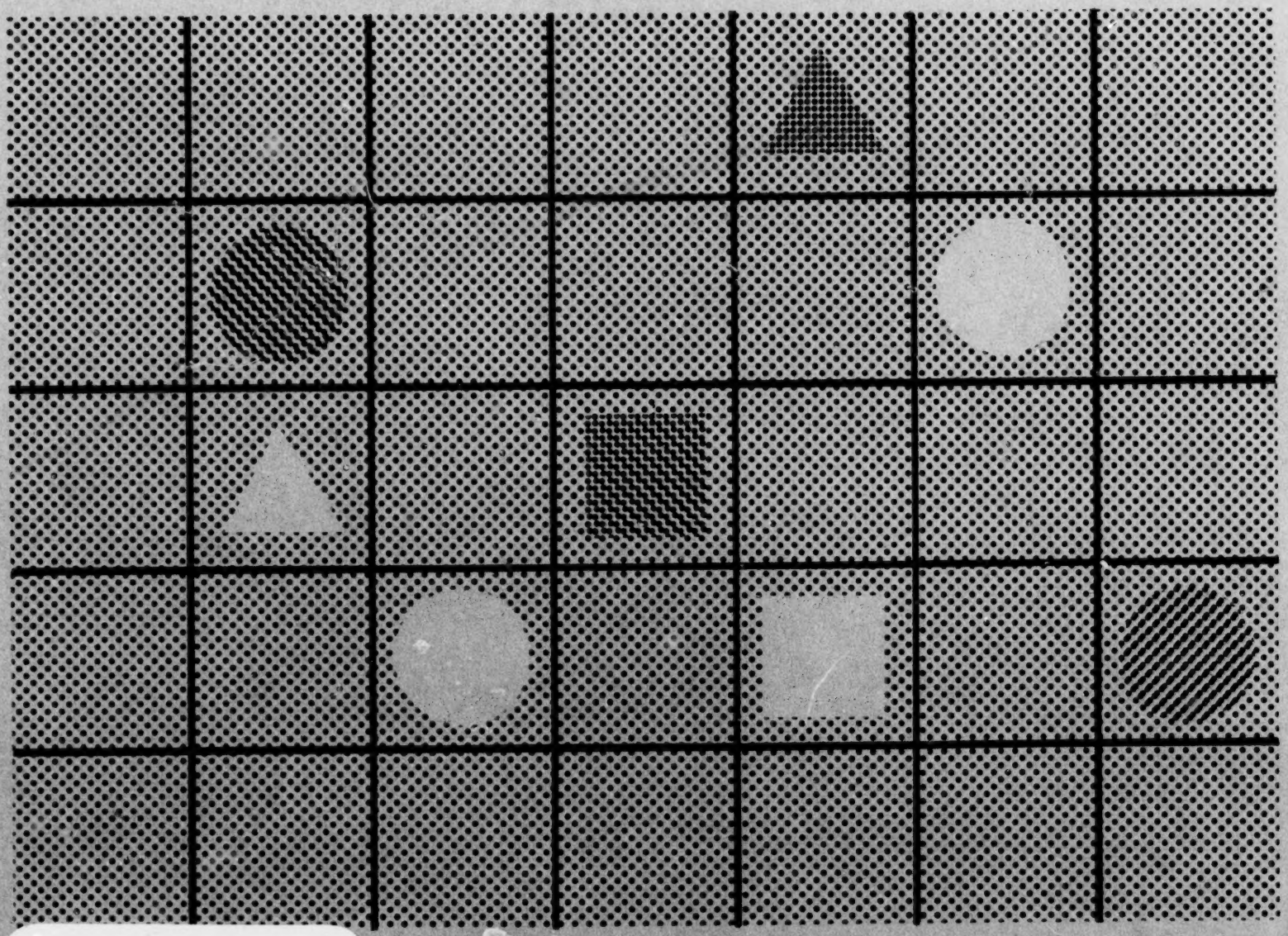
ABSTRACT

The two major purposes of this study are: (1) To provide local school officials with information regarding the performance on basic skills achievement of each student who took the 1970-71 Michigan Educational Assessment Battery, and (2) To provide local officials with information that will assist them in understanding and interpreting their students' scores. The first of the four major sections describes the content of each sub-test (vocabulary, reading, mechanics of written English, and mathematics) in the educational assessment battery. The second describes cautions that must be exercised in the interpretation of individual pupil scores from the program. The third section explains how to interpret the materials that accompany this booklet, and the fourth defines statistical terms used in the educational assessment program and provides technical information regarding the educational assessment battery. (Author/AG)

ED053217

1970-71 INDIVIDUAL PUPIL REPORT: EXPLANATORY MATERIALS

MICHIGAN EDUCATIONAL ASSESSMENT PROGRAM



TM 000 756

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

MICHIGAN DEPARTMENT OF EDUCATION APRIL 1971

State Board of Education

	Term Expires
Edwin L. Novak, O.D., *President* Flint	Jan. 1, 1973
Michael J. Deeb, *Vice President* Detroit	Jan. 1, 1977
Dr. Gorton Riethmiller, *Secretary* Chelsea	Jan. 1, 1975
Thomas J. Brennan, *Treasurer* Dearborn	Jan. 1, 1979
Marilyn Jean Kelly Detroit	Jan. 1, 1977
Annetta Miller Huntington Woods	Jan. 1, 1979
Dr. Charles E. Morton Detroit	Jan. 1, 1973
James F. O'Neil Livonia	Jan. 1, 1975

Dr. John W. Porter, *Superintendent of Public Instruction,*
Chairman, Ex-Officio

William G. Milliken, *Governor*
Member, Ex-Officio

ED053217

FOREWORD

The Michigan Educational Assessment Program was initiated in 1969 by the State Board of Education. During the first year of the Program data concerning the distribution and levels of educational performance of schools and school districts were reported.

This report is the first report of results from the 1970-71 educational assessment. Public Act 38 of 1970 constituted the statutory base for the current educational assessment. The Department, consistent with the requirements of this act, has stated four major objectives of the 1970-71 educational assessment. The four objectives are:

1. to provide members of the State Board of Education, the Executive Office, and the Legislature with information needed for allocating the State's educational resources in a manner best calculated to equalize and improve the quality of educational opportunities for all children of the State;

2. to provide citizens and educators in every school system with basic information regarding their district and its schools;

3. to provide school districts with basic information regarding students that will help students, their parents, and educators to assess their progress; and

4. to provide information regarding the progress of the Michigan educational system as a whole, the progress of its school districts and the progress of its schools over a period of years.

Data regarding the third objective are presented in this report of findings entitled, Individual Pupil Report: Explanatory Materials (Lansing, Michigan: Michigan Department of Education, 1971). Future reports will present information regarding the three other objectives.

Thanks are due to a large number of individuals and groups for making the Michigan Assessment Program a reality: the State Board of Education for proposing it, the Governor and Legislature for actively supporting it, and the Michigan educators for assisting with it. The Program was designed and administered by the Bureau of Research, Evaluation, and Assessment, Michigan Department of Education, with the counsel of several ad hoc advisory groups.

John W. Porter
Superintendent of
Public Instruction

TABLE OF CONTENTS

SECTION

INTRODUCTION 1

1. DESCRIPTION OF THE EDUCATIONAL ASSESSMENT MEASURES 2

Vocabulary 2

Reading 3

Mechanics of Written English 3

Mathematics 3

Composite Achievement 4

2. PRECAUTIONS IN THE INTERPRETATION OF PUPIL SCORES 5

3. INTERPRETATION OF PUPIL SCORES 7

Description of Norming Group 7

Definition of Terms 7

Description of Pupil Roster and Pupil Labels 9

4. STATISTICAL CHARACTERISTICS OF THE 1970-71 MICHIGAN EDUCATIONAL ASSESSMENT BATTERY 12

Definition of Terms 12

Statistical Characteristics of the Fourth Grade Basic Skills Battery 16

Statistical Characteristics of the Seventh Grade Basic Skills Battery 19

Summary. 22

FOOTNOTES 24

INTRODUCTION

This explanatory booklet and the materials that accompany it have two major purposes. This first purpose is to provide local school officials with information regarding the performance on basic skills achievement of each student who took the 1970-71 Michigan Educational Assessment Battery. The second purpose is to provide local officials with information that will assist them in understanding and interpreting their students' scores.

This booklet has four major sections. The first section describes the content of each sub-test - vocabulary, reading, mechanics of written English, and mathematics - in the educational assessment battery. The second section describes cautions that must be exercised in the interpretation of individual pupil scores from the program. The third section explains how to interpret the materials that accompany this booklet. Explanations are provided for pupils scores. The fourth section defines statistical terms used in the educational assessment program and provides technical information regarding the educational assessment battery.

SECTION 1

DESCRIPTION OF THE EDUCATIONAL ASSESSMENT MEASURES

The 1971 assessment battery, like the version given in 1970, was developed according to specifications developed by Educational Testing Service and reviewed by various panels.[1] On the advice of these panels, specific changes were made in approximately 20% of the questions in this year's tests. Some of the changes suggested by panel members had to be dropped in order to retain enough previous test items for comparison of test results between years.

The 1970-71 assessment battery consisted of four different achievement measures--vocabulary, reading, mechanics of written English, and mathematics. Each measure was a separate, timed test at the fourth and seventh grade. It is the purpose of this section to describe these tests.

Vocabulary

The vocabulary test contained 50 verbal analogy problems which were designed to measure students' knowledge of the meaning of words and the relationships between words and concepts. The time allowed to work on this test was 20 minutes at both grades.

---

[1]These panels convened for the purpose of reviewing socioeconomic background, reading, mechanics of written English, and mathematics portions of the battery. The panels consisted of nine teachers, five administrators, four university professors, four content area specialists, two counselors and one school psychologist. Department members and an Educational Testing Service representative were also present at the meetings. The names and affiliations of these panel members are available upon request from the Bureau of Research, Evaluation and Assessment.

Reading

The reading test contained 50 questions which assessed paragraph comprehension, ability to understand words from the context in which they are encountered, and ability to identify the correct synonym for a word. Students at both grade levels were allowed 35 minutes to work on this test.

Mechanics of Written English

The mechanics of written English test consisted of four parts, each separately timed. In part A, spelling, students were to identify misspelled words. The fourth grade test presented 15 items to be completed in five minutes; the seventh grade test presented 20 items to be completed in six minutes. Part B, effectiveness of written expression, contained 14 items for each grade and nine minutes were allowed for its completion. Students were required to select the best way of expressing a thought. Eight minutes were allowed for part C, written usage. Seventh graders were to recognize grammatical errors in 17 items and fourth graders in 14 items. To recognize errors of punctuation and capitalization was the object of part D, punctuation and capitalization. The seventh grade booklet contained 14 items to be completed in seven minutes and the fourth grade booklet had twelve items and allowed eight minutes.

Mathematics

The mathematics test involved mathematical reasoning and problem solving. In addition, problems in the seventh grade test involved

algebraic and geometric concepts. Each grade had 30 minutes in which to answer 40 questions.

Composite Achievement

A composite achievement score was computed for each student. The composite score was obtained by averaging the individual's standard scores on the reading, the mechanics of written English, and the mathematics tests. The test scores were averaged in such a way that each score contributed equally to the average--despite the fact that the number of items was different on the three tests.

It should be noted that the vocabulary test score was not included in the calculation of the composite achievement score. The vocabulary score is believed to respond more slowly to the influence of schooling. Therefore, it was excluded to focus the composite achievement score upon those aspects of basic skills achievement that respond most readily to change.

SECTION 2

PRECAUTIONS IN THE INTERPRETATION OF PUPIL SCORES

A number of precautions should be observed in interpreting the individual pupil scores from the Michigan Educational Assessment Program. The following precautions are particularly important.

All measurement is subject to error. Scores resulting from educational tests are no exception. Therefore, in any use of the results an allowance for error should be made. The discussion on reliability in Section 4 of this report contains estimates of the amount of error associated with each of the assessment instruments on which individual pupil scores are reported.

The assessment results provide a general idea of the basic skills achievement levels of each pupil compared to the basic skill levels of pupils throughout the state. In other words, the tests were purposely made general in their content so that they would be useful in the varied school situations that are to be found in the state. As a result, the particular sample of questions in the test cannot be expected to match perfectly the details of the program in any particular school. Instead, an effort has been made to focus upon the broader outcomes sought by all schools in reading, mechanics of writing, and mathematics. The score of a pupil at any time will be the result of a number of influences. These influences include, but are not limited to: previous educational experience, effectiveness of the school program, teaching effectiveness, home environment, and peer culture. Furthermore, a pupil's score is likely to reflect the combined impact of these forces over several years--even at the fourth grade level.

General achievement batteries are not designed to provide diagnostic information upon which specific instructional plans may be based, though some achievement tests claim to provide so-called "diagnostic item analyses." The Michigan educational assessment tests, being general achievement batteries, do not provide diagnostic information. Rather, they provide a general indication of a pupil's skill level. If a particular student's results on the achievement tests suggest a problem, an appropriate diagnostic instrument may provide information useful for planning instruction to correct the problem. Information about diagnostic tests*in each skill area may be found in the series of Mental Measurements Yearbooks prepared by O.K. Buros. However, the administration of full diagnostic batteries to all pupils is generally unnecessary.

Many areas of pupil achievement and development are not included in the assessment battery. It must be remembered that the goals of local schools are much broader that those assessed by this battery. Therefore persons wishing to judge the overall achievement of pupils must look to additional measures for judgment in other areas of pupil development.

*A test is considered to be diagnostic if it provides a detailed analysis of strengths and weaknesses in an area. Such an analysis will often suggest causes for the deficiencies revealed and indicate possible remedial steps.

SECTION 3

INTERPRETATION OF PUPIL SCORES

The purpose of this section is to provide the reader with information which will facilitate his interpretation of pupil scores. This section contains three parts: a description of the norming group, definitions of the terms necessary for understanding the pupil rosters and labels, and descriptions of the pupil rosters and labels.

Description of the Norming Group

The 1970-71 Michigan Assessment Battery was administered to public school students in the fourth and seventh grades. Excluded were all pupils receiving instruction in special classes for the handicapped in the content areas of reading, English, and mathematics. Those pupils receiving itinerant special education services (e.g., hard of hearing, physically handicapped, educable mentally retarded, emotionally disturbed) and who received instruction in the regular class program in the areas of reading, English and mathematics were included in the assessment. Pupils in non-graded programs who were either identified as fourth or seventh graders or were in their fourth or seventh year beyond kindergarten were included. Also included were shared-time pupils who received instruction in the basic skills in the public school.

Definition of Terms

This part defines the terms which are necessary for understanding the pupil rosters and labels. The defined terms are: mean, standard

deviation, standard scores, and percent below.

Mean

A mean score is an average of a set of scores. It is obtained by adding all of the scores and dividing the sum by the number of scores.

Standard Deviation

In addition to establishing a mean for a distribution of scores, it is often useful to know the "spread" of the scores. Two groups of scores could have the same mean but still be quite different. For example, one district might have children whose scores on composite achievement are very similar and have a mean score of fifty. In this district, the "spread" of scores would be small. Another district might have a number of children with high scores and a number of children with low scores and still have a mean score of fifty. In this district, however, the "spread" of scores would be large.

One common way of indicating the "spread" of a set of scores is to calculate the standard deviation. Usually about two-thirds of the scores will fall between one standard deviation above and one standard deviation below the mean. The larger the standard deviation, the larger will be the "spread" or variability in the scores of a distribution. In the example above, the district with similar scores would have a smaller standard deviation than would the district with the mixture of high and low scores.

Standard Scores

Standard scores are those that are derived from "raw" or response scores using the concepts of mean and standard deviation. In the Michigan Educational Assessment Program, standard scores were developed so that the pupil mean score from any assessment battery measure was fifty and the standard deviation was ten, when computed for all pupils at the same grade level. For example, a pupil with a standard score of forty on reading is one standard deviation below the state mean; a pupil with a score of sixty is one standard deviation above the mean; a pupil with a score of sixty-five is one and one-half standard deviations above the mean; and so forth.

Percent Below

A percent below corresponding to any given standard score is the percentage of pupils in the norm group who received lower scores. Thus, a percent below score of 75 on vocabulary indicates that 75 percent of the pupils in the state received a lower score, and that 25 percent of the pupils in the state received the same or higher score; a percent below score of 16 would mean that 16 percent of the pupils in the state received a lower score and 84 percent of the pupils received the same or a higher score. And so on.

Description of Pupil Roster and Pupil Labels

The pupil roster and pupil labels yield exactly the same information in different formats. Therefore, only the pupil roster will be fully described.

Pupil Roster

Figure 1, on page 10, represents a pupil roster. An example from

ROSTER OF PUPIL SCORES

MICHIGAN DEPARTMENT OF EDUCATION – 1970–71 MICHIGAN EDUCATIONAL ASSESSMENT PROGRAM

School Name ABLE ELEMENTARY SCHOOL (3487) District Name MICHVILLE PUBLIC SCHOOLS (01-010) Grade 4

STUDENT NAME			DATE OF BIRTH		SEX	VOCABULARY			READING			MECHANICS OF WRITTEN ENGLISH			MATH			COMPOSITE ACHIEVEMENT	
Last	First	M.I.	Month	Year		Number Right/Number of Questions	Standard Score	Percent Below	Number Right/Number of Questions	Standard Score	Percent Below	Number Right/Number of Questions	Standard Score	Percent Below	Number Right/Number of Questions	Standard Score	Percent Below	Standard Score	Percent Below
AABERSON	ALEX	M	07	60	M	25/50	53	63	25/50	44	28	35/55	54	61	20/40	48	40	49	43

Figure 1

it will now be described. Alex M. Aaberson was born in July, 1960, is a male, and is in fourth grade. He received a "raw" score of 25 out of a possible 50 questions on the vocabulary test. This "raw" score of 25 translates to a standard score of 53, and 63 percent of the pupils in the state received a lower score. He got 25 of 50 questions correct on reading for a standard score of 44, and 28 percent of the pupils in the state received a lower score. On mechanics of written English, Alex got 35 of 55 questions correct for a standard score of 54, and 61 percent of the students in the state received a lower score. On mathematics, he got 20 of 40 questions correct for a standard score of 48, and 40 percent of the students in the state received a lower score. Finally his composite achievement standard score was 49, and 43 percent of the pupils in the state received a lower score. The remaining names and scores on the roster may be described in a similar manner. A pupil roster is provided for each school in the district which has a fourth and/or seventh grade.

Pupil Labels

The pupil label is displayed below and contains similar information as the pupil roster.

Figure 2

MICHIGAN EDUCATIONAL ASSESSMENT	NAME			GRADE	SEX	DATE OF BIRTH	
	LAST	FIRST	M.I.			MONTH	YEAR
	AABERSON	ALEX	M	4	M	07	60

1970-71	VOCABULARY	READING	MECHANICS OF WRITTEN ENG.	MATHEMATICS	COMPOSITE ACHIEVEMENT
NUMBER RIGHT / QUESTIONS IN TEST	25/50	25/50	35/55	20/40	
STANDARD SCORE	53	44	54	48	49
PERCENT BELOW	63	28	61	40	43

SECTION 4

STATISTICAL CHARACTERISTICS OF THE 1970-71 MICHIGAN EDUCATIONAL ASSESSMENT BATTERY

The purpose of this section is to provide information concerning the statistical properties of the tests. The section is divided into three major parts. The first part will define the statistical terms which are necessary for the readers' understanding of test characteristics. The next part will discuss the statistical characteristics of the achievement measures at the fourth grade and seventh grade levels. The final part will evaluate the test characteristics and will indicate the areas in which additional information is needed.

Definition of Terms

This part defines the statistical terms: validity, reliability, difficulty, speededness, and standard error of measurement.

Validity

The validity of a test is an indication of the extent to which it measures what it is intended to measure. The most important type of validity for achievement tests is content validity. Content validity means that a test which claims to measure elementary mathematics, for example, should contain questions in mathematics and that those questions should be appropriate for the grade level for which the test is intended. The content validity of a test is dependent upon the extent to which the questions in the test constitute a representative sample

of the topics that comprise the subject tested.

It is virtually impossible to experimentally determine the content validity of a test of school achievement, or to report content validity as a numerical coefficient. The best available evidence of validity is found in the test itself and in a description of the procedures used to construct it.

Content validity is likely to be achieved if the development of the tests is the joint responsibility of specialists in test construction and specialists in the skills to be taught. The development of the achievement tests in the Michigan assessment battery has been described in Section 1 and could be expected to produce tests of high content validity.

Concurrent validity is another type of validity. Since it is reported as a correlation coefficient, it is useful in determining the extent to which two tests measure the same subject or characteristic.

Reliability

The reliability of a test provides an estimate of the test's consistency or stability. A test is reliable if it measures consistently whatever it measures; the most reliable tests yield relatively precise results for each student, and a student would receive roughly the same score on such a test if he could take it more than once under the same conditions. The reliability of a test is reported as a coefficient--that is, a two-place decimal figure. A reliability coefficient can range between .00 and 1.00. A test increases in reliability as the coefficient increases in value. An internal consistency method (Kuder-Richardson 20) was used in estimating the reliability of the basic skills tests in the Michigan educational assessment battery. This method employs

information about the length of the test and the extent to which the questions in the test contribute mutually confirming or consistent information.

Difficulty

The difficulty of a test is an indication of how well suited it is to the ability of the group being tested. A test of middle difficulty is appropriate when the group being tested is heterogeneous, such as the group being tested in the Michigan Educational Assessment Program. A test is of middle difficulty when the mean score is near the point midway between a perfect score and the expected chance score. If the test includes 50 items a perfect score would, of course, be 50. The expected chance score, when the 50 items each offer four alternative answers, is 1/4 of 50 or 12.5. Hence the ideal mean (midway between chance and perfect) is 31.25. Expressed as a percentage, this would be 62.5%. In the Michigan educational assessment battery, a percentage significantly higher than 62.5% would indicate an easy test; a percentage significantly lower than 62.5% would indicate a difficult one.

Speededness

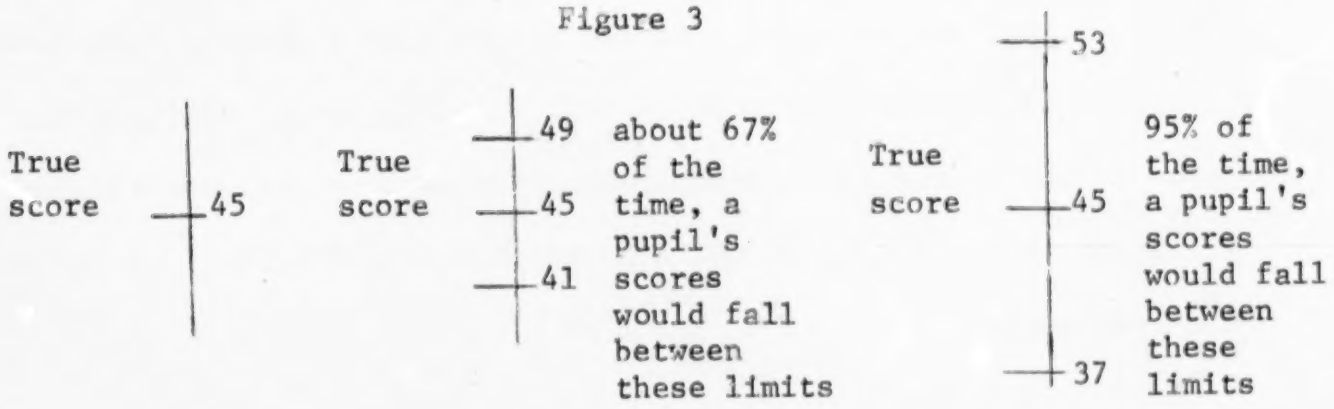
Speededness is a measure of the extent to which test performance is affected by the time limit placed on the test. The criterion used in judging the speededness of a test is two-fold: the proportion of students who answered (1) the last question and (2) the question that is three-quarters of the way through the test. This does not mean that the students answered all the questions up to these points; it means that they reached these particular items. If 80 percent of the students complete the last item and virtually all the students complete 75 percent of the items, the test is usually judged to be unspeeded. It

should be noted that it is desirable for tests in the Michigan Educational Assessment Program to be non-speeded.

Standard Error of Measurement

The standard error of measurement yields an index of error which is likely to occur in a test score. It may be used to form an error band around a score extending from a given number of standard errors of measurement above the score to the same number of standard errors of measurement below the score. Figure 3 illustrates the error band around the true score for a standard error of measurement of 4 score units.

Figure 3



Thus, it can be seen that in the above example the student's true score is 45. However, if the student would take the reading test a large number of times and under the same conditions--with no practice effect or learning taking place--about 67% of the time his score would fall between 41 and 49 (one standard error of measurement), and 95% of the time his true score would fall between 37 and 53 (two standard errors of measurement).

Statistical Characteristics of the Fourth Grade Basic Skills Battery

A representative sample of 1,000 student tests was drawn from the number of fourth grade students who participated in the 1970-71 Michigan Educational Assessment Program. This sample's responses were used to analyze the characteristics of the fourth grade educational assessment battery.

The Test Analysis Sample

The number of questions, means, and standard deviations for the sample and for the entire population are shown in Table 1. From the table it can be seen that the characteristics of the sample conform closely to the characteristics of the entire fourth grade population. Therefore the results presented in the following tables may be generalized to the fourth grade population. Furthermore, information concerning the distribution of raw scores and standard scores are reported in the appendix.

Table 1

STATISTICS ON THE GRADE FOUR 1970-71 EDUCATIONAL ASSESSMENT TESTS FOR THE TEST ANALYSIS SAMPLE AND FOR ALL FOURTH GRADERS TESTED

		Sample		Population	
Score	Number of Questions	Mean	Standard Deviation	Mean	Standard Deviation
Vocabulary	50	21.7	8.4	22.1	8.6
Reading	50	30.6	10.5	31.0	10.5
Mechanics of Written English	55	29.9	10.2	30.7	10.3
Mathematics	40	21.0	7.2	21.8	7.3

Reliability

Table 2 presents the reliability coefficients and standard errors of measurement. The reliability coefficients were: .866 for vocabulary, .925 for reading, .903 for mechanics of written English, .859 for mathematics and .959 for composite achievement. The standard errors of measurement expressed in raw score units were: 3.1 on vocabulary, 2.9 on reading, 3.2 on mechanics of written English, 2.7 on mathematics, and (not defined) on composite achievement. The standard errors of measurement expressed in standard score units were: 3.6 on vocabulary, 2.7 on reading, 3.1 on mechanics of written English, 3.7 on mathematics, and 1.9 on composite achievement. The reliability coefficients are adequately high and the standard errors of measurement are adequately low to justify reporting of individual student scores.

Table 2

RELIABILITY COEFFICIENTS AND STANDARD ERRORS OF MEASUREMENT FOR A SAMPLE OF 1000 FOURTH GRADERS ON THE 1970-71 EDUCATIONAL ASSESSMENT BATTERY

Score	Reliability Coefficient	Standard Error of Measurement	
		Raw Score Units	Standard Score Units
Vocabulary	.866	3.1	3.6
Reading	.925	2.9	2.7
Mechanics of Written English	.903	3.2	3.1
Mathematics	.859	2.7	3.7
Composite Achievement	.959	(not defined)	1.9

Difficulty and Speededness

Table 3 presents the difficulty and speededness indices. The difficulty indices on reading, mechanics of written English, mathematics, and composite achievement are within an acceptable range of middle difficulty. The table does indicate that the vocabulary test was difficult for the fourth grade students. It also indicates that the vocabulary test was speeded.

Table 3

DIFFICULTY AND SPEEDEDNESS FOR A SAMPLE OF 1000 FOURTH GRADERS ON THE 1970-71 EDUCATIONAL ASSESSMENT BATTERY

Score	Average Item Difficulty (Percent passing)	Speededness: Percent reaching last item	Speededness: Percent completing 75% of test
Vocabulary	44.2	63.3	84.2
Reading	62.1	88.0	97.2
Mechanics of Written English	55.8	NA	NA
a. Spelling	NA	90.4	96.5
b. Effectiveness of Written Expression	NA	91.4	96.4
c. Written Usage	NA	92.8	97.2
d. Punctuation and Capitalization	NA	93.6	97.6
Mathematics	54.4	87.0	95.5

Statistical Characteristics of the Seventh Grade Basic Skills Battery

A representative sample of 1,000 student tests was drawn from the number of seventh grade students who participated in the 1970-71 Michigan Educational Assessment Program. This sample's responses were used to analyze the characteristics of the seventh grade educational assessment battery.

The Test Analysis Sample

The number of questions, means, and standard deviations for the sample and for the entire population are shown in Table 4. From the table it can be seen that the characteristics of the sample conform closely to the characteristics of the entire seventh grade population. Therefore the results presented in the following tables may be generalized to the seventh grade population. Furthermore, information concerning the distribution of raw scores and standard scores are reported in the appendix.

Table 4

STATISTICS ON THE GRADE SEVEN 1970-71 EDUCATIONAL ASSESSMENT TESTS FOR THE TEST ANALYSIS SAMPLE AND FOR ALL SEVENTH GRADERS TESTED

Score	Number of Questions	Sample		Population	
		Mean	Standard Deviation	Mean	Standard Deviation
Vocabulary	50	27.2	8.2	27.3	8.3
Reading	50	31.2	9.8	31.2	10.0
Mechanics of Written English	65	33.8	10.1	33.8	10.1
Mathematics	40	21.2	7.6	21.4	7.7

Reliability

Table 5 presents the reliability coefficients and standard errors of measurement. The reliability coefficients were: .866 for vocabulary, .908 for reading, .888 for mechanics of written English, .872 for mathematics and .955 for composite achievement. The standard errors of measurement expressed in raw score units were: 3.0 on vocabulary, 3.0 on reading, 3.4 on mechanics of written English, 2.7 on mathematics, and (not defined) on composite achievement. The standard errors of measurement expressed in standard score units were: 3.7 on vocabulary, 3.0 on reading, 3.4 on mechanics of written English, 3.5 on mathematics, and 1.0 on composite achievement. The reliability coefficients are adequately high and the standard errors of measurement are adequately low to justify reporting of individual student scores.

Table 5

RELIABILITY COEFFICIENTS AND STANDARD ERRORS OF MEASUREMENT FOR A SAMPLE OF 1000 SEVENTH GRADERS ON THE 1970-71 EDUCATIONAL ASSESSMENT BATTERY

Score	Reliability Coefficient	Standard Error of Measurement: Raw Score Units	Standard Error of Measurement: Standard Score Units
Vocabulary	.866	3.0	3.7
Reading	.908	3.0	3.0
Mechanics of Written English	.888	3.4	3.4
Mathematics	.872	2.7	3.5
Composite Achievement	.955	(not defined)	1.9

Difficulty and Speededness

Table 6 presents the difficulty and speededness indices. The difficulty indices on vocabulary, reading, mechanics of written English, mathematics, and composite achievement are within an acceptable range of middle difficulty.

Table 6

DIFFICULTY AND SPEEDEDNESS FOR A SAMPLE OF 1000 SEVENTH GRADERS ON THE 1970-71 EDUCATIONAL ASSESSMENT BATTERY

Score	Average Item Difficulty (Percent passing)	Speededness: Percent reaching last item	Speededness: Percent completing 75% of test
Vocabulary	54.6	82.7	95.5
Reading	62.3	90.4	98.4
Mechanics of Written English	52.0	NA	NA
a. Spelling	NA	92.5	98.1
b. Effectiveness of Written Expression	NA	97.1	98.7
c. Written Usage	NA	95.5	97.8
d. Punctuation and Capitalization	NA	94.3	98.2
Mathematics	53.6	86.8	96.5

Summary

A perusal of tables 1 through 6 indicates the statistical strength of the achievement batteries. However, certain types of statistical data are presently not available in the Michigan Educational Assessment Program.

The Michigan Educational Assessment Program has provided reasonable content validity in the achievement battery. Programs are currently underway to make the content of the assessment battery even more relevant to Michigan's curricula. At present, concurrent validity information is lacking. However, collection of this information is anticipated in the future.

APPENDIX

Table of Vocabulary Raw Scores, Standard Scores and Percents Below

Raw Score	Grade 4 Standard Score	Grade 4 Number of Pupils	Grade 4 Percent Below	Grade 7 Standard Score	Grade 7 Number of Pupils	Grade 7 Percent Below
50	83	5	100.0	77	5	100.0
49	81	12	100.0	76	39	100.0
48	80	41	100.0	75	102	99.9
47	79	83	99.9	74	303	99.7
46	78	150	99.8	73	441	99.4
45	77	276	99.6	71	766	99.0
44	76	407	99.4	70	1152	98.2
43	74	592	99.0	69	1492	97.3
42	73	798	98.5	68	1993	96.0
41	72	1024	97.9	67	2485	94.5
40	71	1343	97.0	65	2901	92.7
39	70	1536	96.0	64	3644	90.4
38	69	1813	94.9	63	4018	87.9
37	67	2100	93.6	62	4533	85.0
36	66	2485	92.0	61	4937	81.9
35	65	2765	90.3	59	5386	78.5
34	64	3019	88.4	58	5744	74.9
33	63	3355	86.3	57	5871	71.2
32	62	3533	84.0	56	6075	67.4
31	60	3926	81.6	54	6497	63.3
30	59	4134	78.9	53	6607	59.2
29	58	4387	76.2	52	6672	55.0
28	57	4660	73.2	51	6454	50.9
27	56	4889	70.2	50	6606	46.8
26	55	5256	66.9	48	6502	42.7
25	53	5500	63.4	47	6484	38.6
24	52	5799	59.7	46	6331	34.6
23	51	6040	55.9	45	6293	30.7
22	50	6315	51.9	44	6172	26.8
21	49	6744	47.7	42	5992	23.1
20	48	7021	43.3	41	5812	19.4
19	46	7265	38.7	40	5426	16.0
18	45	7458	34.0	39	4940	12.9
17	44	7524	29.2	38	4362	10.1
16	43	7588	24.5	36	3881	7.7
15	42	7117	20.0	35	3260	5.7
14	41	6495	15.9	34	2550	4.1
13	39	5838	12.2	33	1931	2.8
12	38	4922	9.1	31	1435	1.9
11	37	3857	6.7	30	1043	1.3
10	36	3201	4.6	29	722	0.8
9	35	2350	3.2	28	483	0.5
8	34	1769	2.0	27	311	0.3
7	32	1117	1.3	25	212	0.2
6	31	838	0.8	24	122	0.1
5	30	516	0.5	23	54	0.1
4	29	364	0.3	22	60	0.1
3	28	202	0.1	21	44	0.0
2	27	135	0.0	19	28	0.0
1	25	54	0.0	18	13	0.0
0	24	25	0.0	17	9	0.0

Table of Reading Raw Scores, Standard Scores and Percents Below

Raw Score	Grade 4 Standard Score	Grade 4 Number of Pupils	Grade 4 Percent Below	Grade 7 Standard Score	Grade 7 Number of Pupils	Grade 7 Percent Below
50	68	70	100.0	69	318	99.8
49	67	278	99.8	68	988	99.2
48	66	758	99.3	67	1784	98.1
47	65	1663	98.3	66	2567	96.4
46	64	2948	96.4	65	3349	94.3
45	63	4104	93.8	64	4009	91.8
44	62	5350	90.4	63	4525	89.0
43	61	6243	86.5	62	4882	85.9
42	60	6674	82.3	61	5204	82.7
41	60	6732	78.1	60	5535	79.2
40	59	6837	73.8	59	5706	75.6
39	58	6544	69.6	58	5720	72.0
38	57	6368	65.6	57	5687	68.4
37	56	5985	61.9	56	5725	64.8
36	55	5751	58.2	55	5662	61.3
35	54	5568	54.7	54	5531	57.8
34	53	5232	51.4	53	5573	54.3
33	52	4977	48.3	52	5208	51.0
32	51	4612	45.4	51	5051	47.9
31	50	4372	42.6	50	5029	44.7
30	49	4264	40.0	49	4825	41.7
29	48	3967	37.5	48	4778	38.7
28	47	3838	35.0	47	4567	35.8
27	46	3697	32.7	46	4397	33.1
26	45	3520	30.5	45	4383	30.3
25	44	3347	28.4	44	4127	27.7
24	43	3308	26.3	43	4048	25.2
23	42	3155	24.3	42	4082	22.6
22	41	3071	22.4	41	3918	20.1
21	40	3130	20.4	40	3833	17.7
20	39	3087	18.5	39	3596	15.5
19	38	3172	16.5	38	3331	13.4
18	38	3067	14.5	37	3282	11.3
17	37	3266	12.5	36	3236	9.3
16	36	3244	10.4	35	2827	7.5
15	35	3100	8.5	34	2637	5.9
14	34	2950	6.6	33	2339	4.4
13	33	2602	5.0	32	1993	3.1
12	32	2192	3.6	31	1579	2.1
11	31	1814	2.5	30	1187	1.4
10	30	1286	1.7	29	822	0.9
9	29	928	1.1	28	550	0.5
8	28	682	0.6	27	317	0.3
7	27	372	0.4	26	210	0.2
6	26	222	0.3	25	124	0.1
5	25	127	0.2	24	64	0.1
4	24	108	0.1	23	46	0.1
3	23	70	0.1	22	38	0.0
2	22	63	0.0	21	30	0.0
1	21	42	0.0	20	22	0.0
0	20	16	0.0	19	13	0.0

Table of Mechanics of Written English Raw Scores, Standard Scores and Percents Below

	Grade 4			Grade 7		
Raw Scores	Standard Score	Number of Pupils	Percent Below	Standard Score	Number of Pupils	Percent Below
65				81	1	100.0
64				80	4	100.0
63				79	13	100.0
62				78	47	100.0
61				77	74	99.9
60				76	160	99.8
59				75	223	99.7
58				74	338	99.5
57				73	434	99.2
56				72	616	98.8
55	74	45	100.0	71	781	98.3
54	73	127	99.9	70	996	97.7
53	72	298	99.7	69	1195	96.9
52	71	544	99.4	68	1458	96.0
51	70	784	98.9	67	1682	95.0
50	69	1115	98.2	66	1943	93.7
49	68	1532	97.2	65	2309	92.3
48	67	1975	96.0	64	2580	90.7
47	66	2426	94.4	63	3019	88.8
46	65	2837	92.6	62	3292	86.7
45	64	3385	90.5	61	3678	84.4
44	63	3645	88.2	60	3918	81.9
43	62	3958	85.7	59	4218	79.3
42	61	4264	83.0	58	4526	76.4
41	60	4509	80.2	57	4864	73.4
40	59	4781	77.2	56	5046	70.2
39	58	4816	74.1	55	5451	66.8
38	57	5047	71.0	54	5478	63.4
37	56	5110	67.7	53	5696	59.8
36	55	5047	64.6	52	5692	56.2
35	54	5061	61.4	51	5764	52.6
34	53	5003	58.2	50	5880	48.9
33	52	5089	55.0	49	5749	45.3
32	51	5134	51.8	48	5631	41.8
31	50	5054	48.6	47	5473	38.3
30	49	4983	45.4	46	5392	34.9
29	48	4890	42.4	45	5305	31.6
28	47	4841	39.3	44	5112	28.4
27	46	4888	36.2	43	4779	25.4
26	45	4746	33.2	42	4645	22.5
25	44	4735	30.2	41	4407	19.7
24	44	4669	27.3	40	3989	17.2
23	43	4418	24.5	39	3874	14.8
22	42	4453	21.7	38	3745	12.4
21	41	4232	19.0	37	3339	10.3
20	40	4202	16.4	36	3078	8.4
19	39	3960	13.9	35	2621	6.7
18	38	3560	11.7	34	2356	5.3
17	37	3271	9.6	33	2021	4.0
16	36	3008	7.7	32	1564	3.0
15	35	2718	6.0	31	1396	2.1
14	34	2285	4.6	30	973	1.5
13	33	1965	3.3	29	734	1.1
12	32	1526	2.4	29	543	0.7
11	31	1143	1.6	28	359	0.5
10	30	878	1.1	27	260	0.3
9	29	642	0.7	26	160	0.2
8	28	416	0.4	25	97	0.2
7	27	237	0.3	24	73	0.1
6	26	145	0.2	23	52	0.1
5	25	98	0.1	22	47	0.1
4	24	66	0.1	21	35	0.0
3	23	43	0.0	20	24	0.0
2	22	32	0.0	19	23	0.0
1	21	17	0.0	18	9	0.0
0	20	12	0.0	17	6	0.0

Table of Mathematics Raw Scores, Standard Scores and Percents Below

	Grade 4			Grade 7		
Raw Score	Standard Score	Number of Pupils	Percent Below	Standard Score	Number of Pupils	Percent Below
40	75	77	100.0	74	55	100.0
39	74	233	99.8	73	257	99.8
38	72	587	99.4	71	644	99.4
37	71	972	98.8	70	1141	98.7
36	70	1551	97.8	69	1841	97.5
35	68	2176	96.5	68	2592	95.9
34	67	2867	94.7	66	3350	93.8
33	65	3454	92.5	65	3925	91.3
32	64	4139	89.9	64	4538	88.4
31	63	4902	86.8	62	4805	85.4
30	61	5571	83.2	61	5086	82.2
29	60	5911	79.5	60	5425	78.8
28	59	6396	75.5	59	5748	75.2
27	57	6624	71.3	57	6054	71.4
26	56	6897	66.9	56	6135	67.5
25	54	7096	62.5	55	6195	63.6
24	53	7018	58.0	53	6379	59.6
23	52	7231	53.5	52	6565	55.4
22	50	7114	49.0	51	6677	51.2
21	49	7158	44.4	49	6715	47.0
20	48	7058	40.0	48	6958	42.6
19	46	7098	35.5	47	6858	38.3
18	45	7078	31.0	46	6703	34.1
17	43	6925	26.7	44	6825	29.8
16	42	6722	22.4	43	6561	25.7
15	41	6325	18.4	42	6402	21.6
14	39	6129	14.6	40	6258	17.7
13	38	5495	11.1	39	5756	14.0
12	37	4653	8.1	38	5226	10.8
11	35	3888	5.7	36	4682	7.8
10	34	3052	3.8	35	3889	5.4
9	32	2294	2.3	34	2898	3.5
8	31	1499	1.4	33	2161	2.2
7	30	934	0.8	31	1567	1.2
6	28	577	0.4	30	918	0.6
5	27	302	0.2	29	504	0.3
4	26	167	0.1	27	253	0.1
3	24	97	0.1	26	116	0.1
2	23	47	0.0	25	50	0.0
1	22	28	0.0	24	31	0.0
0	20	7	0.0	22	9	0.0

Table of Composite Achievement Standard Scores and Percents Below

Standard Score	4th Grade Number of Pupils	4th Grade Percent Below	7th Grade Number of Pupils	7th Grade Percent Below
74	0	100.0	4	100.0
73	0	100.0	13	100.0
72	5	100.0	64	99.9
71	41	100.0	172	99.8
70	134	99.9	402	99.6
69	342	99.7	707	99.1
68	637	99.3	1079	98.4
67	1107	98.6	1444	97.5
66	1644	97.5	1917	96.3
65	2367	96.0	2377	94.8
64	3085	94.0	2892	92.9
63	3748	91.7	3376	90.8
62	4344	88.9	3902	88.3
61	4942	85.8	4196	85.6
60	5321	82.4	4641	82.7
59	5662	78.8	4920	79.5
58	5893	75.0	5179	76.2
57	5981	71.2	4354	72.8
56	5925	67.5	5550	69.3
55	5931	63.7	5664	65.7
54	5776	60.0	5873	61.9
53	5752	56.4	5908	58.2
52	5553	52.8	5828	54.5
51	5376	49.4	5797	50.8
50	5173	46.1	5694	47.1
49	5258	42.8	5755	43.5
48	5093	39.5	5603	39.9
47	5006	36.3	5487	36.4
46	4872	33.2	5393	33.0
45	4721	30.2	5081	29.7
44	4591	27.3	5030	26.5
43	4619	24.4	4831	23.4
42	4501	21.5	4633	20.5
41	4382	18.7	4478	17.6
40	4347	16.0	4189	15.0
39	4072	13.4	3989	12.4
38	3980	10.9	3839	10.0
37	3711	8.5	3446	7.8
36	3410	6.3	3070	5.8
35	2883	4.5	2702	4.1
34	2372	3.0	2201	2.7
33	1805	1.8	1765	1.6
32	1250	1.0	1131	0.9
31	746	0.6	698	0.4
30	400	0.3	351	0.2
29	213	0.2	158	0.1
28	118	0.1	72	0.0
27	72	0.1	38	0.0
26	31	0.0	17	0.0
25	29	0.0	5	0.0
24	17	0.0	4	0.0
23	6	0.0	2	0.0
22	1	0.0	2	0.0

Raw and Standard Score Means and Standard Deviations

Test	Statistic	Grade 4 Raw Score	Grade 4 Standard Score	Grade 7 Raw Score	Grade 7 Standard Score
Vocabulary	Mean	22.1	50.0	27.3	50.0
	Standard Deviation	8.6	10.0	8.3	10.0
Reading	Mean	31.0	50.0	31.2	50.2
	Standard Deviation	10.5	10.0	10.0	10.0
Mechanics of Written English	Mean	30.7	50.0	33.8	49.8
	Standard Deviation	10.3	9.9	10.1	10.1
Mathematics	Mean	21.8	50.0	21.4	50.0
	Standard Deviation	7.3	10.0	7.7	10.0
Composite Achievement	Mean	n.a.	50.0	n.a.	50.1
	Standard Deviation	n.a.	9.2	n.a.	9.1